Modelos SARIMA (Auto-Regressivo Integrado de Médias Móveis com Sazonalidade): parte 1

Prof. Caio Azevedo

Introdução

- Vamos retornar ao processos ARIMA(p,d,q). Com efeito, seja

$$Y_t = Y_{t-1} + \epsilon_t, t = 1, 2, ...,$$

em que $Y_0 \equiv c$. Então, temos que:

$$Y_t = c + \epsilon_1 + \epsilon_2 + ... + \epsilon_t.$$

- Portanto $E[Y_t] = c$ e $\mathcal{V}[Y_t] = (t-1)\sigma^2$, $\forall t$. Isto é, trata-se de um processo não estacionário na variância.

Introdução

- Considere agora o processo:

$$Y_t = Y_{t-1} + (1 + \theta B)\,\epsilon_t = Y_{t-1} + \omega_t, \omega_t \sim MA(1). \qquad (1)$$

- Do processo acima temos que $\mathcal{E}(Y_t) = \mathcal{E}(Y_{t-1})$ e, assumindo que $Y_0 \equiv c$, temos que processo é estacionário em termos da média. Isto é, não há "tendência genuína".
- No entanto, a perturbação (ω_t), em geral, não é um ruído branco. Então $\{Y_t\}$ tende a apresentar valores, por longos períodos acima e por longos períodos abaixo da média.

Introdução

- O fenômeno anterior é conhecido como tendência estocástica.
- Note que o processo (1) pode ser reescrito como:

$$(1 - B)Y_t = (1 + \theta B)\epsilon_t,$$

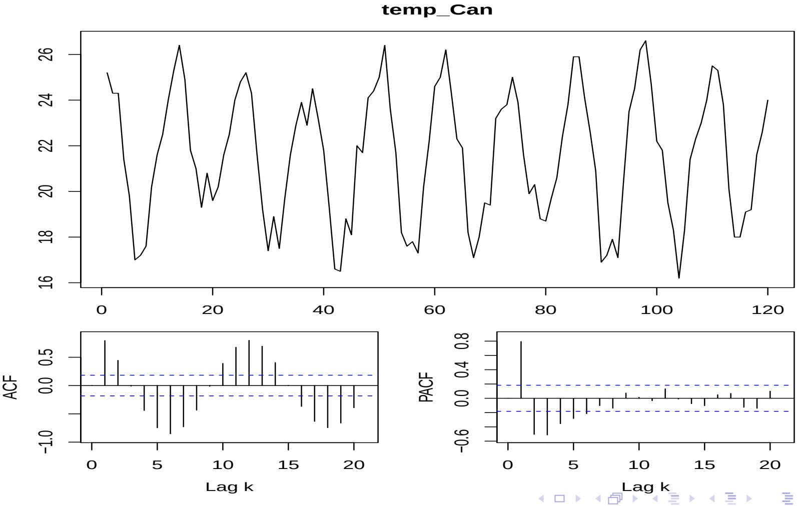
  o qual é membro da família ARIMA(p,q,d).

- Por outro lado, muitas séries exibem comportamentos sazonais. A seguir discutiremos como modificar o processo ARIMA (ARMA) para que capture sazonalidade estocástica.

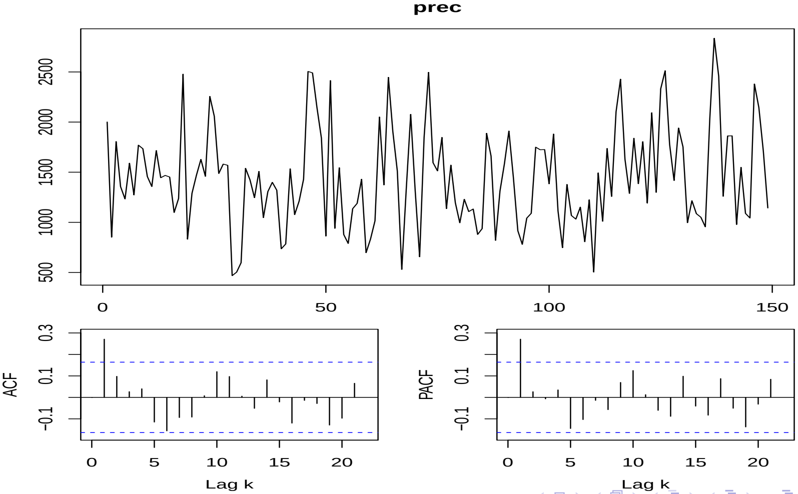
Introdução

- No exemplo das temperaturas médias mensais de Cananéia, notamos que há uma similaridade entre as temperaturas nos mesmos meses, ao logo dos diferentes anos.
- Ou seja, ao logo dos "Janeiros", "Fevereiros" etc, ao longo dos anos. Isto é, observa-se um comportamento periódico em tempos múltiplos de $s = 12$.
- Então seria conveniente ter um modelo que explique a ST (também) nos tempos múltiplos de s.

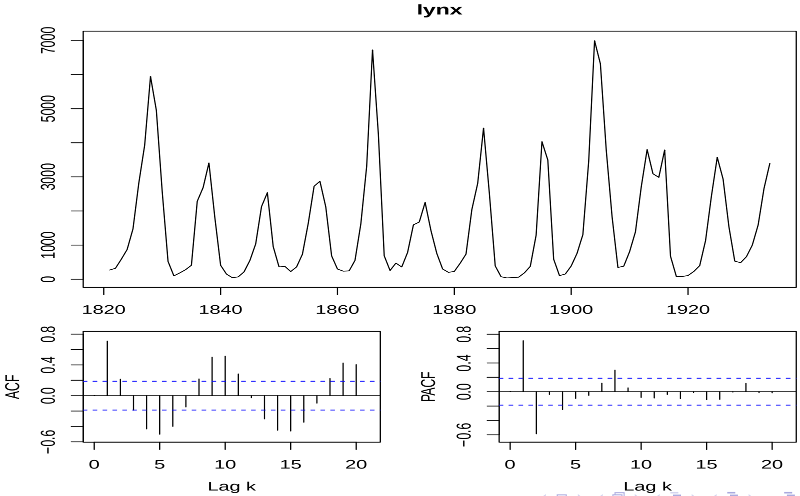
Temperaturas médias mensais de Cananéia



Precipitação pluviométrica de Fortaleza



Número de linces presos



Modelos SARMA (puros)

- Os modelos Sazonais Autoregressivos de Média Móveis **puros**, $SARMA(P,Q)_s$ são definidos como:

$$
\begin{aligned}
\Phi(B^s)(Y_t - \mu) &= \Theta(B^s)\epsilon_t \\
\Phi(B^s) &= 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \Phi_3 B^{3s} - \ldots - \Phi_P B^{Ps} \\
\Theta(B^s) &= 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \Theta_3 B^{3s} + \ldots + \Theta_P B^{Ps} \\
\epsilon_t &\sim RB(0, \sigma^2).
\end{aligned}
$$

Modelos SARMA (puros)

- Quando $s = 1$, temos um modelo ARMA(p,q).
- Exemplo, se $P = Q = 1$ e $s = 2$, temos que:

$$
\begin{aligned}
(1 - \Phi_1 B^2)(Y_t - \mu) &= (1 + \Theta_1 B^2)\epsilon_t \\
Y_t &= \mu + \Phi_1(Y_{t-2} - \mu) + \Theta_1 \epsilon_{t-2} + \epsilon_t.
\end{aligned}
$$

- O termo "sazonalidade pura" indica que somente as observações/ruídos brancos de ordem s (sazonalidade), são considerados no modelo. Em geral, (somente) valores da FAC dessa ordem ($\rho(k)$) tendem a ser (mais) significativos.

Modelos SARMA

- (Cont.) Ou seja, o modelo anterior se **assemelha** à um modelo ARMA. Contudo, as defasagens nas observações (Y_t) e nos erros (ϵ_t), que são consideradas no modelo, dependem da ordem da sazonalidade (s).
- À rigor, os modelos SARMA, tem uma estrutura ARMA, em múltiplos como função da ordem da sazonalidade (s). Ou seja, tal processo é como um processo ARMA, porém com operadores definidos em múltiplos de s.

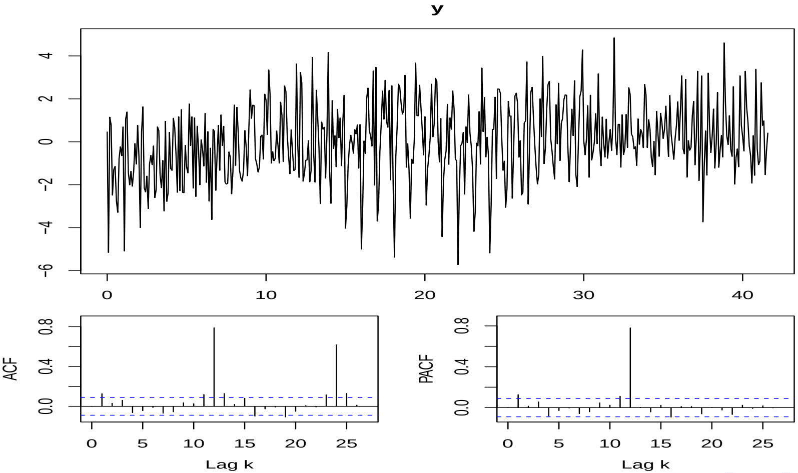
Modelos SARMA

- Assim, o que fora visto em termos de estimação, previsão e diagnóstico (aqui, aqui) para os modelos ARMA, pode ser adaptado para os modelos SARMA.
- Portanto, o processo SARMA é causal somente quando as raízes do polinômio $\Phi(z^s)$ são maiores que um, em modulo, e é invertível somente quando as raízes do polinômio $\Theta(z^s)$ são maiores que um, em modulo.

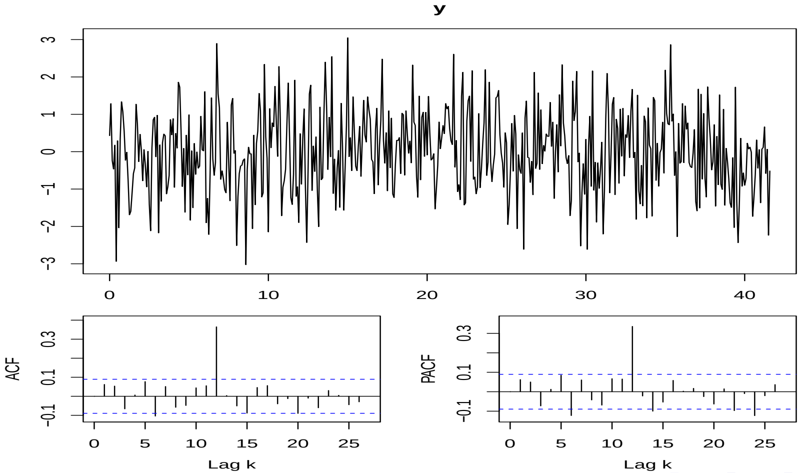
Modelos SARMA

- Além disso, como pode ser visto nos gráficos das respectivas séries simuladas (próximos slides), as correlações tendem a ser maiores (significativas) para observações com distância equivalente aos períodos de sazonalidade (s). Ou seja, $\rho(k \times s), k = 1, 2, ...$.
- Portanto, esse tipo de modelo é apropriado quando as correlações (significativas) são sazonais (ou seja, que oscilam/há uma periodicidade) mas não-seriais (que tende a decrescer com o aumento da distância entre as observações).

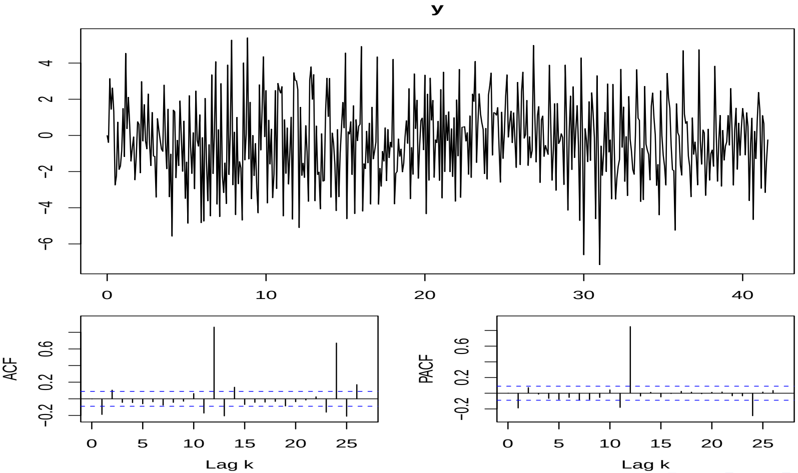
Série simulada $SARMA(1,0)_{12}, \Phi_1 = 0,8$



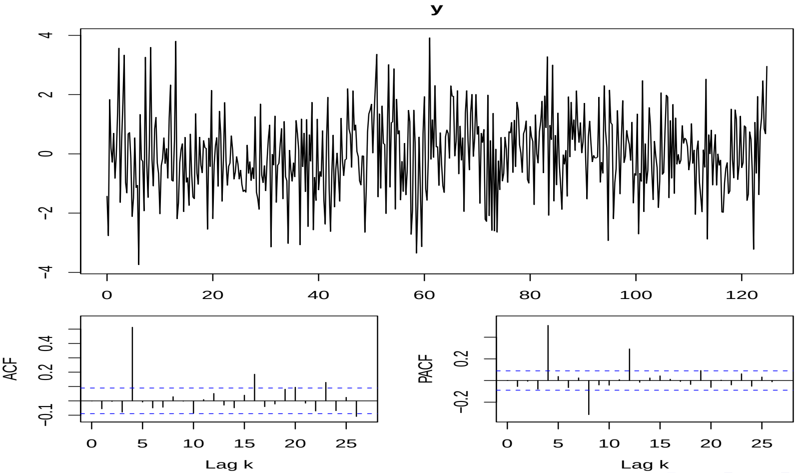
Série simulada $SARMA(0,1)_{12}, \Theta_1 = 0,5$



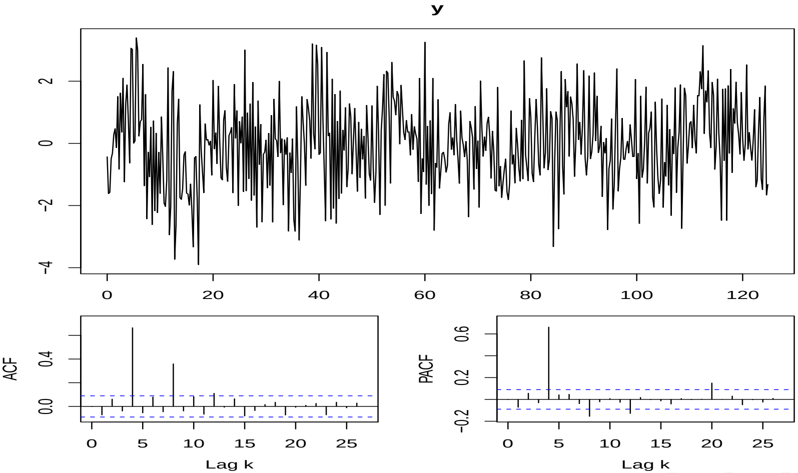
Série simulada $SARMA(1,1)_{12}, \Phi_1 = 0,8; \Theta_1 = 0,5$



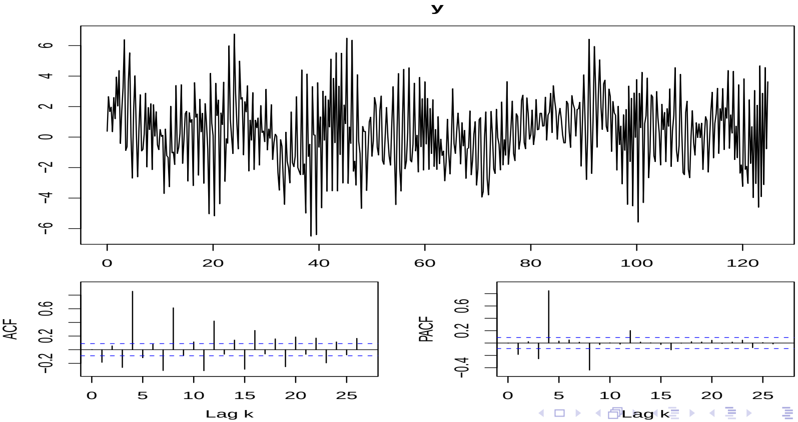
Série simulada $SARMA(3,0)_4, \mathbf{\Phi} = (0,8; -0,5; 0,3)'$



Série simulada $SARMA(0,3)_4, \Theta = (0,7; 0,5; 0,3)$



Série simulada $SARMA(3,3)_4$, $\mathbf{\Phi} = (0,8; -0,5; 0,3)'$, $\mathbf{\Theta} = (0,7; 0,5; 0,3)'$



Mais exemplos: $SARMA(1,0)_{12}$

- Tal modelo é dado por:

$$\left(1 - \Phi_1 B^{12}\right) Y_t = \epsilon_t,$$

$$Y_t = \Phi_1 Y_{t-12} + \epsilon_t.$$

- Note que este modelo é causal e invertível se $|\Phi_1| < 1$. Analogamente aos processos ARMA, a FACV é dada por:

$$\gamma(k) = \begin{cases} \sigma^2/(1-\Phi_1^2), & \text{se } k = 0 \\ \sigma^2 \Phi_1^{k/12}/(1-\Phi_1^2), & \text{se } k = ..., -24, -12, 12, 24 \\ 0, \text{c.c.} \end{cases}$$

Mais exemplos: $SARMA(1,0)_{12}$

- Observe que para as defasagens que não são múltiplos de 12, as autocovariâncias são iguais a zero.
- O modelo anterior seria apropriado, por exemplo, no caso em que as observações correspondam a vendas mensais, e, assim, as vendas de uma dado mês dependem das vendas do mesmo mês, no ano anterior, segundo uma estrutura AR(1).
- Exercício: apresente a versão do modelo anterior com $\mu \neq 0$ e calcule a FAC: $\rho(k)$.

Mais exemplos: $SARMA(0,1)_4$

- Tal modelo é dado por:

$$Y_t = \left(1 - \Theta_1 B^4\right) \epsilon_t$$

$$Y_t = \Theta_1 \epsilon_{t-4} + \epsilon_t.$$

- Note que este modelo é invertível se $|\Theta_1| < 1$. A FACV é dada por:

$$\gamma(k) = \begin{cases} \sigma^2(1+\Theta_1^2), & \text{se } k = 0 \\ \Theta_1 \sigma^2, & \text{se } |k| = 4 \\ 0, \text{c.c.} \end{cases}$$

Modelos SARMA

- A FACV (FAC) dos processos $SARMA(P,Q)_s$ apresenta o mesmo comportamento (qualitativo) que os processos ARMA, nos tempos múltiplos de s.
- O modelo anterior seria apropriado, por exemplo, em que as observações correspondem a vendas mensais, e, assim as vendas de uma dado mês dependem das vendas do mesmo mês, no ano anterior, segunda uma estrutura MA(1).

Modelos SARMA

- Exercício: apresente a versão do modelo anterior com $\mu \neq 0$ e calcule a respectiva FAC: $\rho(k)$.
- Exercício: Encontre as condições de causalidade e invertibilidade do processo $SARMA(1,1)_{12}$ definido como: $\left(1 - \Phi_1 B^{12}\right) Y_t = \left(1 + \Theta_1 B^{12}\right) \epsilon_t$

Comentários

- Notamos, assim, que os modelos *SARMA* são como modelos *ARMA*, com estruturas de correlação particulares induzidas pela sazonalidade do processo.
- A FAC e a FACP dos processos SARMA são semelhantes as do processo ARMA, definida pela estrutura de sazonalidade.
- Entretanto, comumente, na prática, encontramos ST sazonais não estacionárias. Por exemplo, os dados sobre o número de passageiros (próximo slide).

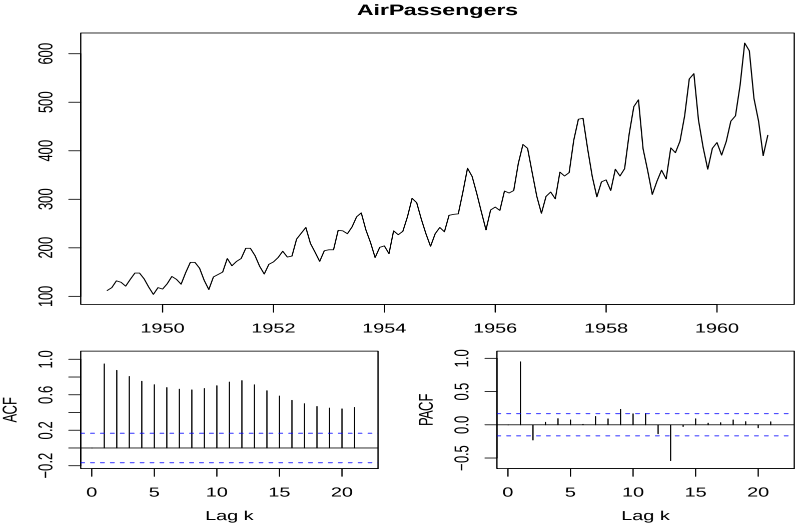


Figura: Totais mensais de passageiros de companhias aéreas internacionais, 1949 a 1960.

Modelos SARIMA (puros)

- Os modelos SARMA são úteis para séries estacionárias com sazonalidade pura.
- Quando a ST apresenta sazonalidade pura, mas não estacionariedade, podemos considerar uma versão ARIMA, para os modelos SARMA.
- Os modelo SARIMA é construído a partir do modelo SARMA, de modo semelhante aos modelos ARIMA e ARMA.
- Primeiramente, definamos o operador diferença sazonal.

$$\nabla_s^D Y_t = (1 - B^s)^D \, Y_t, D \in \{1, 2, ...\}$$

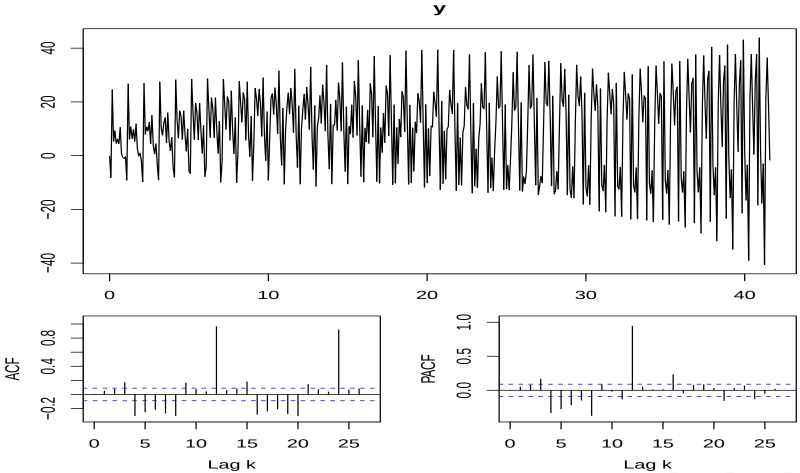
Modelos SARIMA (puros)

- O modelo $SARIMA(P,D,Q)_s$ (puro, no mesmo sentido do que fora apresentado para o os modelos SARMA) é definido como

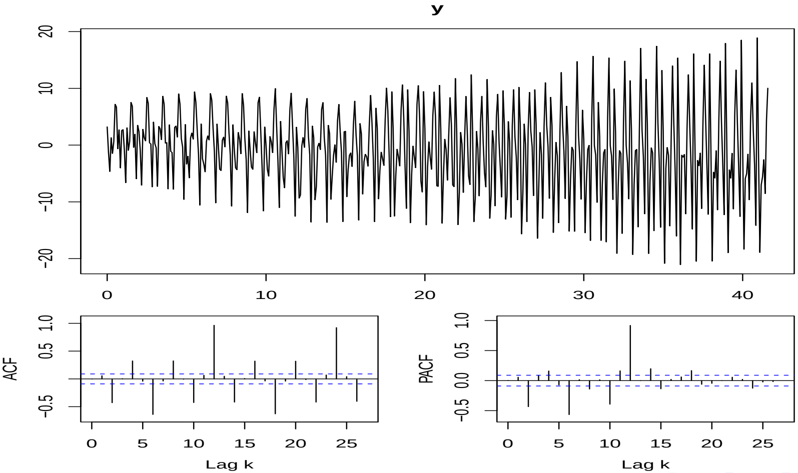
$$\Phi(B^s)\left(1-B^s\right)^D\left(Y_t-\mu\right)=\Theta\left(B^s\right)\epsilon_t$$

- Note que o processo acima é como o ARIMA, com periodicidade s.
- Com efeito, se $s=1$, temos um processo ARIMA.
- Assim, o que fora visto em termos de estimação, previsão e diagnóstico (aqui) para os modelos ARMA, pode ser adaptado para os modelos SARMA.
- Nos próximos slides veremos alguns exemplos gráficos de processos SARIMA.

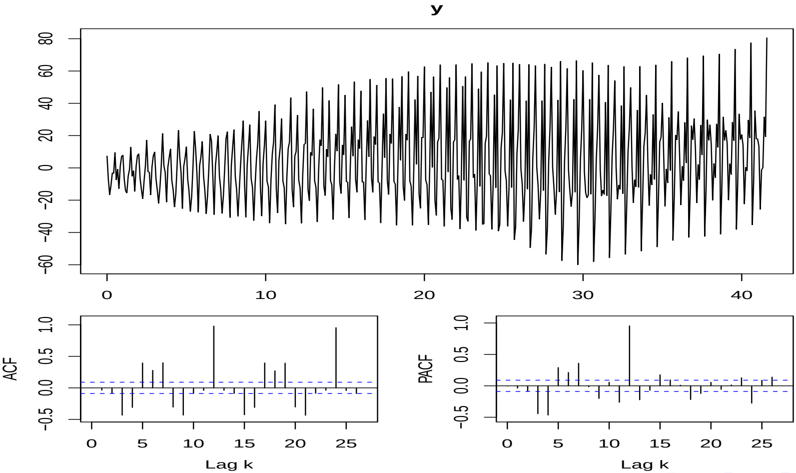
Série simulada $SARIMA(1,1,0)_{12}, \Phi_1 = 0,8$



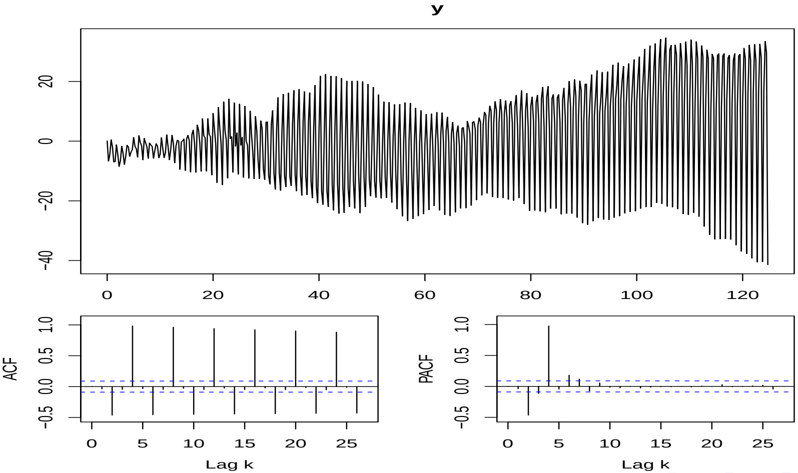
Série simulada $SARIMA(0,1,1)_{12}, \Theta_1 = 0,5$



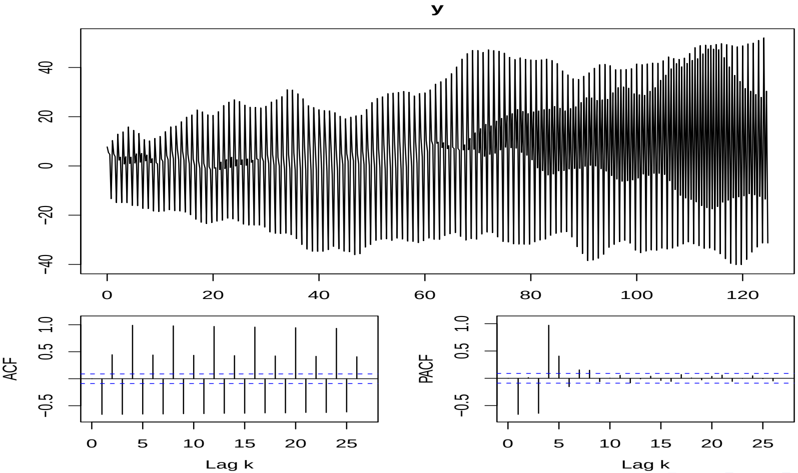
Série simulada $SARIMA(1,1,1)_{12}, \Phi_1 = 0,8, \Theta_1 = 0,5$



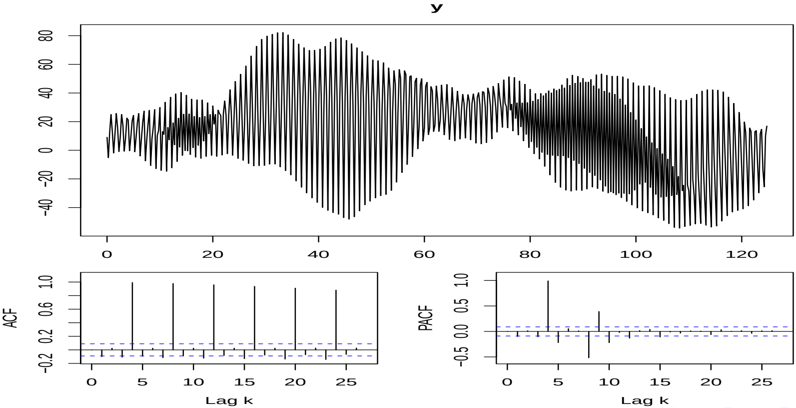
Série simulada $SARIMA(3,1,0)_4$, $\Phi = (0,8;-0,5;0,3)'$



Série simulada $SARIMA(0,1,3)_4, \boldsymbol{\Theta} = (0,7;0,5;0,3)$



Série simulada $SARIMA(3,1,3)_4$, $\mathbf{\Phi} = (0,8; -0,5; 0,3)'$, $\mathbf{\Theta} = (0,7; 0,5; 0,3)'$



Comentários

- Notamos assim, que os modelos *SARIMA* são como modelos *ARIMA*, com estruturas de correlação particulares, induzidas pela sazonalidade do processo.
- A FAC e a FACP dos processos SARIMA são semelhantes as do processo ARIMA, definida pela estrutura de sazonalidade.
- Note que as estruturas (amostrais) da FAC e FACP para os modelos SARIMA (SARMA) podem não indicar, de forma "simples", possíveis modelos apropriados.

Comentários

- Além disso, esses processos são úteis quando há uma sazonalidade "pura", ou seja, quando no processo faz sentido considerar somente os termos correspondentes à sazonalidade (termos de ordem s).
- Entretanto, em alguma situações, além da dependência de termos relativos à sazonalidade (ordem s), pode haver dependência de termos antes ou depois de s (sazonalidade multiplicativa).
- Por exemplo, a temperatura/pluviosidade de uma dado mês pode depender não somente das medidas do mesmo mês de anos anteriores, mas também de outros meses.

Comentários

- Com efeito, é muito mais comum encontrar sazonalidade multiplicativa, do que pura.
- Para esse tipo de situação é mais apropriado combinar a parte sazonal pura com a parte ARIMA (ARMA).
- Neste caso, temos uma classe ampla de modelos para ST:
  - Os Modelos Multiplicativos Sazonais Autoregressivos Integrados de Médias Moveis: $SARIMA(p, d, q)(P, D, Q)_s$.
- Note que devemos distinguir os tipos de modelo SARIMA, através da colocação ou não dos termos $(p, d, q), (P, D, Q)$.

SARIMA (multplicativo)

- Dizemos que $\{Y_t\}$ é um processo SARIMA$(p,d,q)(P,D,Q)_s$ de média μ se for definido como:

$$\Phi(B^s)\phi(B)(1-B^s)^D(1-B)^d(Y_t-\mu) = \Theta(B^s)\theta(B)\epsilon_t,$$

em que $\epsilon_t \sim RB(0,\sigma^2)$ e

$$\begin{aligned}
\Phi(B^s) &= 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps}, \\
\Theta(B^s) &= 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \cdots + \Theta_Q B^{Qs}, \\
\phi(B) &= 1 - \phi_1 B^1 - \phi_2 B^2 - \cdots - \phi_p B^p, \\
\theta(B) &= 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q.
\end{aligned}$$

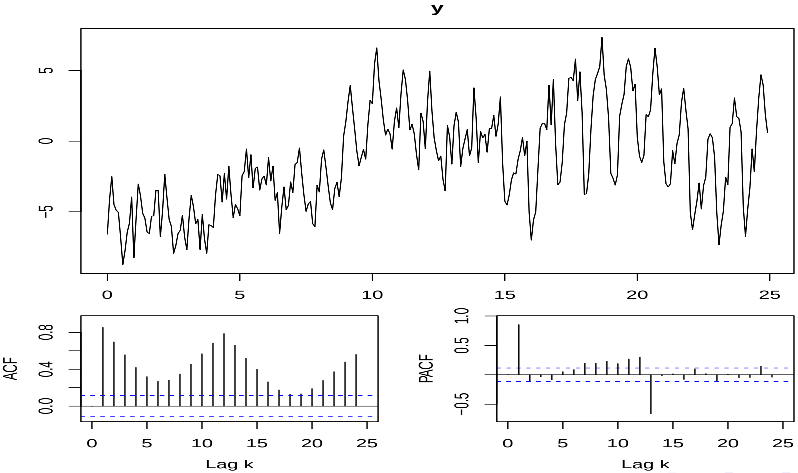
SARIMA

- A análise da estrutura de dependência desse modelo, por exemplo através da FAC e FACP, realiza-se combinando o comportamento da parte sazonal (periódica) e da parte pura (lag a lag). Veja os exemplos a seguir.
- Com relação à estimação, a partir da estrutura do slide anterior, podemos, por exemplo, construir a verossimilhança (ainda que de forma aproximada) baseada em todas as observações.

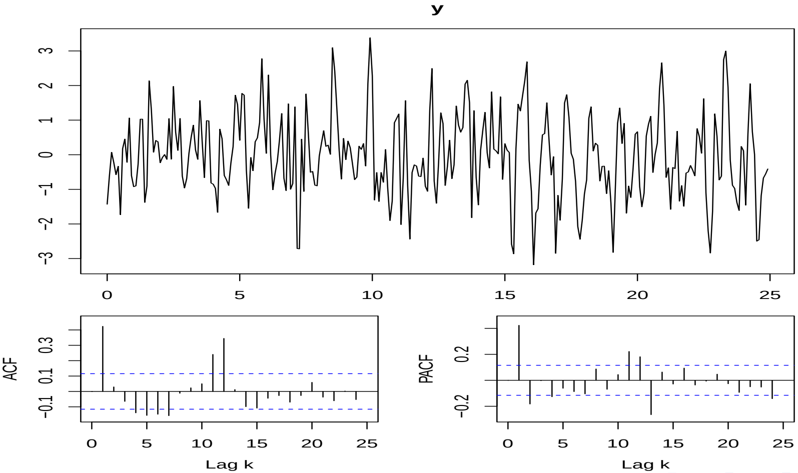
SARIMA

- O que fora apresentado (inclusive para outros métodos de estimação) para os modelos ARMA e ARIMA, podem ser adaptados. Veja também aqui e aqui.
- Com relação à parte residual, podemos utilizar o que fora apresentado aqui. Falaremos sobre previsão, mais a frente.

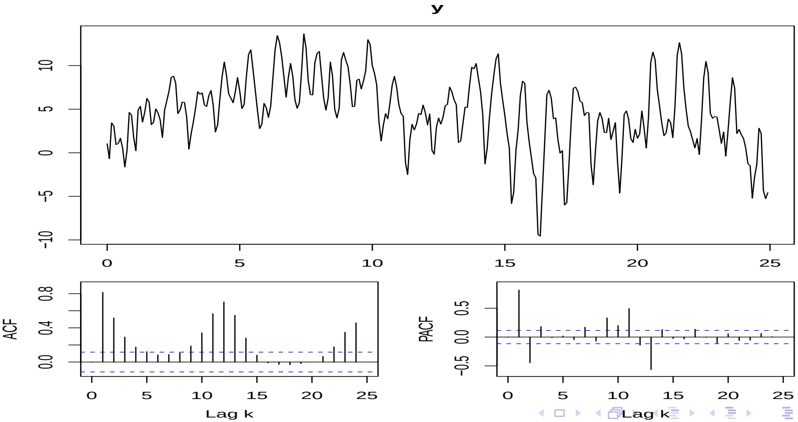
Série simulada $SARMA(1,0)(1,0)_{12}, \phi_1, \Phi_1 = 0,8$



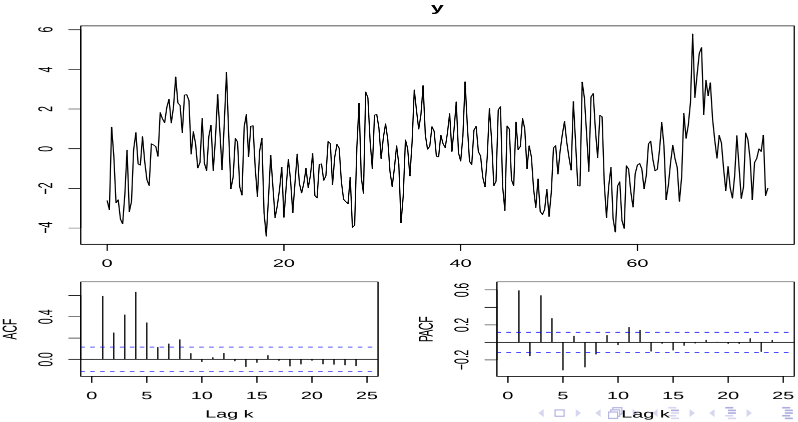
Série simulada $SARMA(0,1)(0,1)_{12}, \theta_1, \Theta_1 = 0,5$



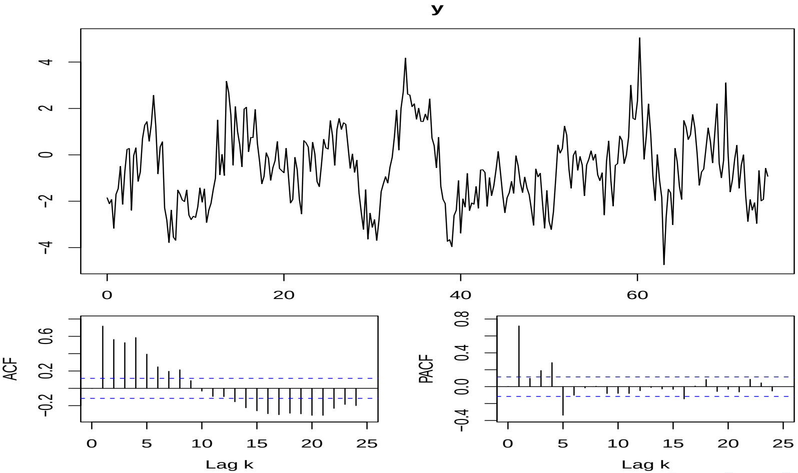
Série simulada $SARMA(1,1)(1,1)_{12}$, $\phi_1, \Phi_1 = 0,8$; $\theta_1, \Theta_1 = 0,5$



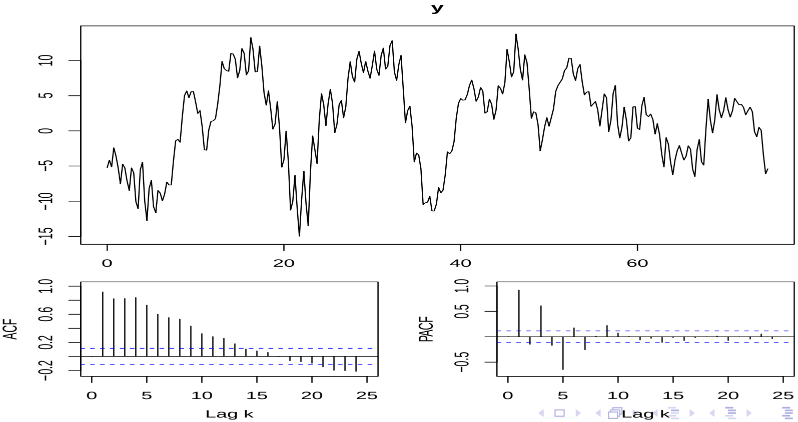
Série simulada $SARMA(3,0)(3,0)_4$, $\phi, \Phi = (0,8; -0,5; 0,3)'$



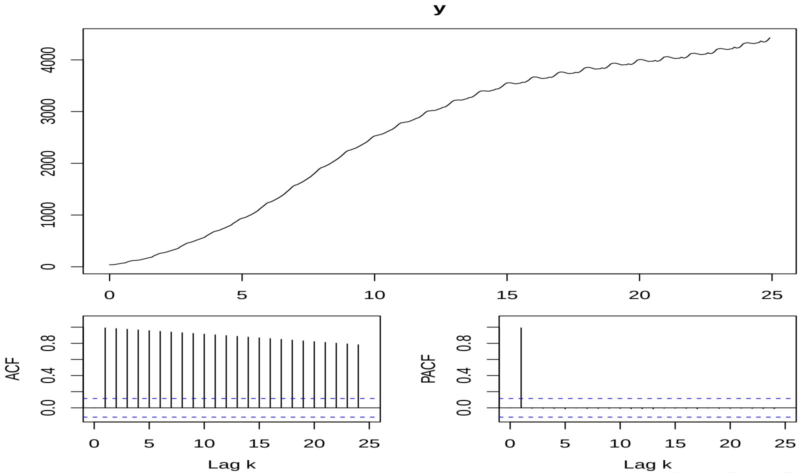
Série simulada $SARMA(0,3)(0,3)_4$, $\boldsymbol{\theta}, \boldsymbol{\Theta} = (0,7; 0,5; 0,3)'$



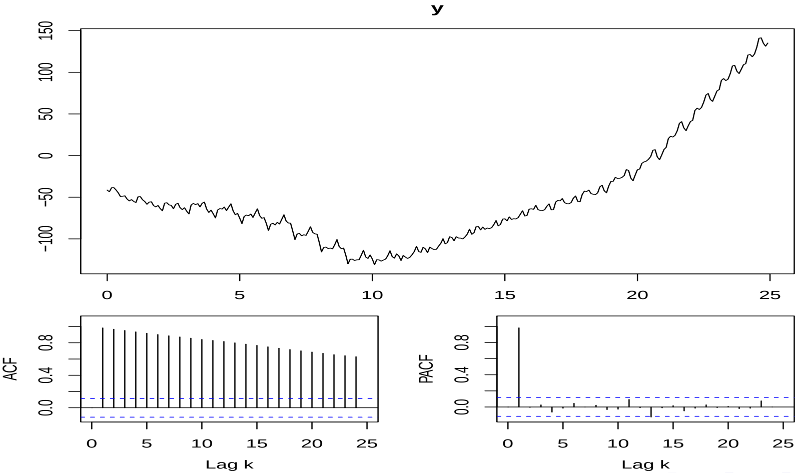
Série simulada $SARMA(3,3)(3,3)_4$, $\boldsymbol{\phi}, \boldsymbol{\Phi} = (0,8; -0,5; 0,3)'$, $\boldsymbol{\theta}, \boldsymbol{\Theta} = (0,7; 0,5; 0,3)'$



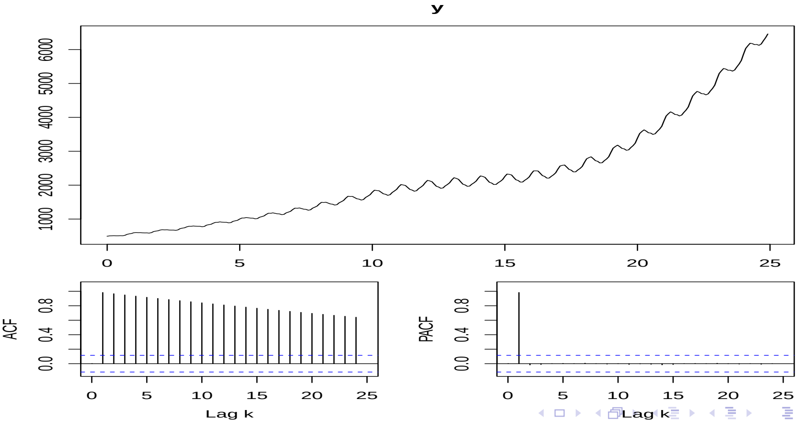
Série simulada $SARIMA(1,1,0)(1,1,0)_{12}, \phi_1, \Phi_1 = 0,8$



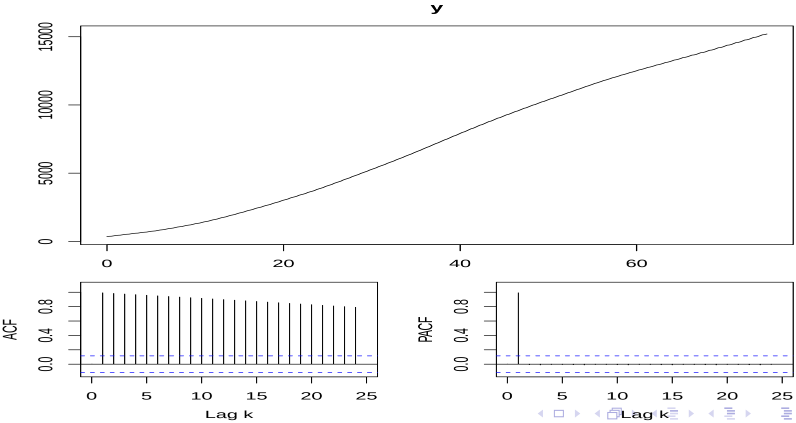
Série simulada $SARIMA(0,1,1)(0,1,1)_{12}, \theta_1, \Theta_1 = 0,5$



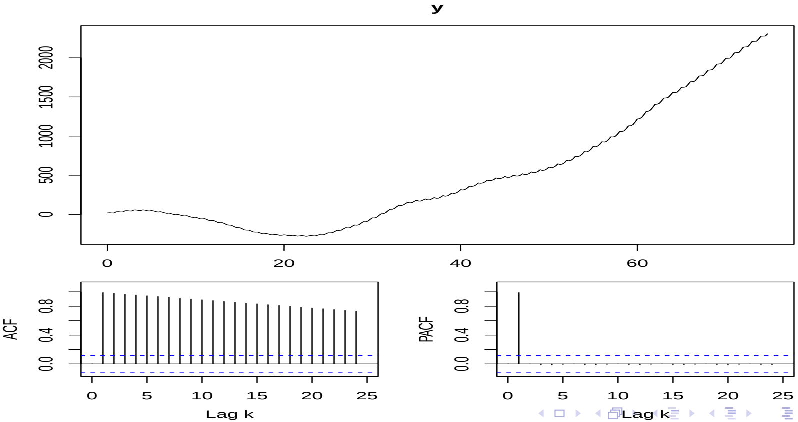
Série simulada $SARIMA(1,1,1)(1,1,1)_{12}$, $\phi_1, \Phi_1 = 0,8$, $\theta_1, \Theta_1 = 0,5$



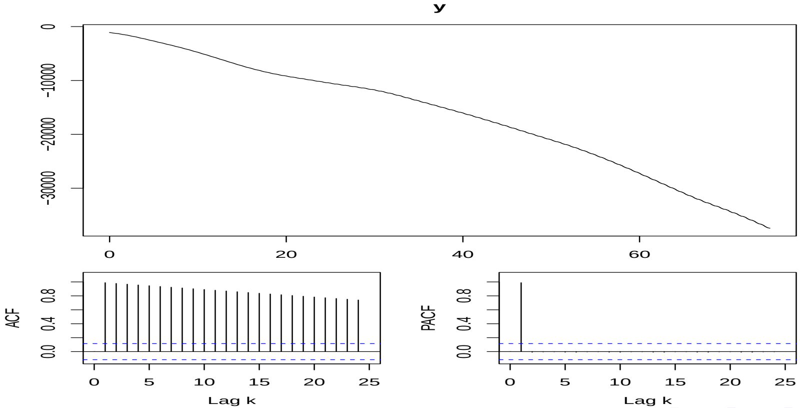
Série simulada $SARIMA(3,1,0)(3,1,0)_4$, $\phi, \Phi = (0,8; -0,5; 0,3)'$



Série simulada $SARIMA(0,1,3)(0,1,3)_4$, $\theta, \Theta = (0,7;0,5;0,3)$



Série simulada $SARIMA(3,1,3)(3,1,3)_4$, $\phi, \Phi = (0,8; -0,5; 0,3)'$, $\theta, \Theta = (0,7; 0,5; 0,3)'$



SARMA

- Um caso particular dos modelos SARIMA são os processos SARIMA não-integrados ou SARMA, nos quais não se considera a diferenciação.
- Como visto anteriormente, dizemos que $\{Y_t\}$ é um processo SARMA$(p,q)(P,Q)_s$ de média μ se for definido como:

$$\Phi(B^s)\phi(B)(Y_t - \mu) = \Theta(B^s)\theta(B)\epsilon_t,$$

em que $\epsilon_t \sim RB(0, \sigma^2)$ e as outras componentes são como definidas anteriormente.

SARMA

- Um processo SARMA pode ser reescrito como um processo ARMA, como veremos a seguir.
- **Exemplo:** considere $\{Y_t\}$ um processo SARMA$(1,0)(1,0)_{12}$, o qual é escrito como:

$$\begin{aligned}
(1-\phi B)(1-\Phi B^{12})Y_t &= \epsilon_t, \\
(1-\phi B-\Phi B^{12}+\phi\Phi B^{13})Y_t &= \epsilon_t,
\end{aligned}$$

que nada mais é do que um ARMA$(13,0)$ com $\phi_1=\phi$, $\phi_2=\phi_3=\cdots=\phi_{11}=0$, $\phi_{12}=\Phi$ e $\phi_{13}=\phi\Phi$.

SARMA

- No caso geral temos que um modelo SARMA$(p, q)(P, Q)_s$ é um modelo ARMA$(p + sP, q + sQ)$ com restrições (resultado particularmente útil nas questões relativas a estimação e previsão).
- Por outro lado, o modelo SARMA é mais interpretável, parcimonioso e não traz restrições relativas a seus parâmetros.

Alguns casos particulares (lista não exaustiva)

- $SARIMA(p, d, q)(P, D, Q)_{s=1} = SARIMA(p, d, q)$.
- $SARIMA(p, d = 0, q)(P, D = 0, Q)_s = SARMA(p, q)(P, Q)_s$.
- $SARIMA(p, d = 0, q)(P, D = 0, Q)_{s=1} = ARMA(p, q)$.
- $SARIMA(p = 0, d = 0, q = 0)(P, D, Q)_s = SARIMA(P, D, Q)_s$.
- $SARMA(p = 0, q = 0)(P, Q)_s = SARMA(P, Q)_s$.